



Rethinking the problem of misinformation and its solutions

new media & society
1–25

© The Author(s) 2026

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14614448261428635

journals.sagepub.com/home/nms



Sacha Altay 

University of Zurich, Switzerland

Abstract

The way a problem is framed shapes its solutions. This article reframes the problem of misinformation and examines the implications of this shift for interventions against misinformation. It advances five arguments that challenge common narratives about misinformation and invite us to rethink both the problem and its solutions. For instance, exposure to misinformation is lower than often believed, people are less gullible than commonly assumed, and misinformation often reflects, rather than causes, underlying sociopolitical issues. These insights point toward strategies that address the root causes of the problem rather than surface symptoms. Key shifts include focusing on the demand for misinformation, fostering trust in reliable sources, and strengthening democratic institutions. Combating misinformation effectively requires a clear understanding of the problem and a break with popular misconceptions about it.

Keywords

Fact-checking, interventions, media literacy, misinformation, news, trust

Introduction

People across the world are worried about misinformation (Newman et al., 2024), scientists are studying it more closely than ever before (Allen et al., 2020), and policymakers are looking for solutions to curb its spread and limit its impact. Following the 2016 US presidential election, efforts to combat misinformation have intensified. Fact-checking initiatives have expanded, various kinds of media literacy training have been developed

Corresponding author:

Sacha Altay, Department of Political Science, University of Zurich, Zürich, 8050, Switzerland.

Email: sacha.yesilaltay@uzh.ch

to help people navigate the information ecosystem, and numerous psychological interventions have been tested to deter the sharing of misinformation. In parallel, our scientific understanding of misinformation and the information ecosystem has improved.

In this article, I highlight a disconnect between this growing body of knowledge and the premises on which interventions against misinformation are built—specifically the idea that misinformation is widespread and a major cause of complex societal issues, that people are gullible and easily swayed by misinformation, and that they share and believe misinformation because they *lack* certain attributes, such as literacy skills or cognitive abilities. These ideas have gained traction in recent years, fueled by headlines on the influence of fake news on the 2016 US presidential election and the Brexit vote (Anderson, 2021) and the surge of false information during the COVID-19 pandemic (the “infodemic,” Simon and Camargo, 2021). However, these fears have proved largely exaggerated (Altay et al., 2023a; Budak et al., 2024; Nyhan, 2020), and few attempts have been made to move away from these narratives and build interventions on stronger, more evidence-based foundations. This is not to say that current interventions are ineffective or that researchers are unaware of these challenges (Allen et al., 2025; Guay et al., 2023a; Modirrousta-Galian and Higham, 2023). Rather, this article is intended as a contribution to an ongoing, collective, and reflexive endeavor aimed at enhancing these interventions. Its goal is to propose new ways to think about the problem of misinformation and how to best address it.

The article is organized as follows. First, I present past reviews on interventions against misinformation and introduce a theoretical framework. Second, I develop five arguments about online misinformation and outline their implications for interventions against misinformation (see Table 1 for a summary). The arguments are representative of the growing body of knowledge on misinformation. Third, I review the most common interventions against misinformation and highlight their strengths and limitations. I detail the extent to which the theoretical foundations of these interventions are in line with our current knowledge of the information environment. Fourth, I propose concrete ways to adapt and hopefully improve interventions against misinformation. At the end of the article, I question whether the objective of interventions against misinformation, most notably reducing misperceptions, could be achieved more efficiently by other means, such as by increasing the acceptance of reliable information. I also stress the need for systemic solutions and the importance of targeting not only the visible aspects of the problem (e.g. false beliefs) but also its underlying causes (e.g. polarization or widespread mistrust).

Ultimately, this article offers new ways of thinking about the problem of misinformation and how best to address it by proposing to (i) shift some of the effort currently devoted to combating misinformation toward increasing the acceptance and consumption of reliable information; (ii) target the handful of powerful individuals who give misinformation most of its visibility; (iii) move beyond false information to address problematic and harmful content more broadly, such as hateful, deceptive, or misleading information; (iv) look beyond individual-level frames and consider the sociopolitical factors that enable the spread of problematic content and explore ways to strengthen norms and institutions that can mitigate its impact.

Literature review

Misinformation is used in this article as an umbrella term encompassing false or misleading content shared without necessarily the intent to harm (Altay et al., 2023b). It is distinct from disinformation, which is spread with the intent to cause harm (Bennett and Livingston, 2018). Because offline misinformation is hard to study, most empirical work focuses on online misinformation in the form of text or URLs. Specifically, research has focused on news fact-checked as false on topics related to politics and health (Pfänder and Altay, 2025). However, we will see that these definitions are not without flaws, that false information is not necessarily harmful, and that true information can be used in misleading ways.

Interventions against misinformation have been under scrutiny recently. A systematic review found that most interventions focused on fact-checking, lacked theoretical foundations, and did not address the causes of misinformation susceptibility, such as motivated reasoning or partisanship (Ziemer and Rothmund, 2024). Most interventions focus on surface-level features and do not address the reasons why people look for, engage with, and believe misinformation. There is a need to move beyond individual-level interventions toward more ambitious system-level solutions, such as platform design changes, tech regulations, or support for free and independent media (Roozenbeek et al., 2023). A scoping review recently highlighted the need for interventions to adapt to video-based social media, to be tested in ecological settings outside the lab, to target people older than 65 years, and to extend their scope outside the United States and Western Europe (Gwiażdźński et al., 2023). Few interventions have been tested in Global South countries (Badrinathan, 2021; Badrinathan and Chauchard, 2024; Blair et al., 2024), and much of the misinformation literature—including the articles cited in this article—focuses on Global North countries, which is a major limitation.

When it comes to measuring the effectiveness of interventions, the goal is no longer solely to reduce belief in false news but also to increase belief in true news (Acerbi et al., 2022; Altay et al., 2024a; Guay et al., 2023b). Researchers increasingly recognize the potential side effects of these interventions, such as increasing distrust in true news (Modirrousta-Galian and Higham, 2023). They caution against relying excessively on truth discernment and emphasize the importance of considering other outcomes such as trust, policy support, and behaviors (Spampatti, 2025).

Most interventions have been tested in experimental conditions that do not resemble online environments and inflate the effectiveness of interventions (Roozenbeek et al., 2024). In experimental settings, all participants are exposed to the treatment (e.g. media literacy tips), the target content (e.g. false headlines), and the outcomes are tested shortly after exposure to the treatment. Yet, for an intervention to be effective beyond experimental settings, several conditions must be met. Importantly, individuals must be in an environment where the intervention can be deployed (e.g. they need to use social media) and must encounter unwanted content (e.g. be exposed to misinformation)—Roozenbeek et al. (2024) offer an overview of such obstacles. Drawing on implementation science, I apply the framework of efficacy versus effectiveness (Roozenbeek et al., 2024). Efficacy refers to the effects of interventions under ideal, controlled conditions, whereas effectiveness refers to their effects outside experimental settings (Singal et al., 2014).

This framework is particularly relevant for interventions against misinformation because few studies have tested their *effectiveness* (e.g. Lin et al., 2024; Nyhan and Reifler, 2015) and have largely focused on testing their *efficacy* in controlled experimental settings (e.g. Arechar et al., 2023; Guess et al., 2020a).

Five arguments about online misinformation

Exposure to misinformation is low

The Internet is flooded with content, most of which does not adhere to strict journalistic or scientific standards. However, to understand the influence of online misinformation, one must go beyond its mere existence and consider its reach. Like all cultural products, the success of online content follows a power law: a small fraction of content receives most of the attention, while the rest goes largely unnoticed (Acerbi, 2020). Content matters only if it is consumed. Fake news has no impact if no one sees it. Rather than asking how common misinformation is (probably very), we need to understand how often people encounter it and in which context.

Measuring misinformation exposure is notoriously difficult because defining and identifying misinformation at scale is challenging. To date, the most exhaustive estimates define misinformation at the source level and combine web browsing data of hundreds of thousands of individuals with reliability ratings of thousands of websites and major social media accounts. Most sources are classified as unreliable (and included in the misinformation category) because of their low journalistic standards of credibility and transparency, not because they repeatedly or predominantly publish false content (Altay et al., 2022). These studies found that in Europe and the United States, misinformation represents between 0.7% and 6% of people's online news diet (Allen et al., 2020; Altay et al., 2022; Cordonier and Brest, 2021; Grinberg et al., 2019; Guess et al., 2019; Osmundsen et al., 2021). This proportion is higher on platforms like Facebook, where in some countries it represents up to one-fifth of the news people are exposed to or engage with (Allen et al., 2021b; Altay et al., 2022; Guess et al., 2021a). Misinformation is assumed to be rampant on messaging apps like WhatsApp, yet early evidence suggests that its prevalence is limited and comparable to that on social media (Chauchard and Garimella, 2022). One reason is that people do not use social media and the web primarily for news or to stay informed (Newman et al., 2025), but for entertainment, socializing, shopping, or gaming. Only about 5% of the time people spend online is dedicated to news consumption (Cordonier and Brest, 2021). Because people are exposed to little news and civic content online (Nyhan et al., 2023), and most misinformation is related to news events, politics, or society, only about 0.15% of people's online media diet is composed of misinformation (Allen et al., 2020; Cordonier and Brest, 2021). Yet, these are averages, and the distribution of misinformation consumption is not normal. A small number of people account for most of the misinformation consumed and shared online (Grinberg et al., 2019), mirroring the concentration of toxic comments or online conflicts (Robertson et al., 2024). A sizable portion of the population avoids not just false news but also reliable news (Allen et al., 2020; Cordonier and Brest, 2021; Wojcieszak et al., 2023).

Because false news consumption is low in Western democracies, training people to spot it, debunk it, or discourage its sharing is bound to have modest effects outside experimental settings. It has been estimated that interventions wiping out all misinformation would only be as effective at improving the overall accuracy of people's beliefs as interventions increasing the acceptance of reliable information by 1% (Acerbi et al., 2022). And since people encounter true news much more often than false news, interventions teaching people to recognize biases and manipulation techniques in the news will primarily affect reliable news. For instance, the critical thinking skills acquired through media literacy training are primarily applied to scrutinize true news. While a certain dose of skepticism and vigilance is certainly healthy, the evidence presented in the section below suggests that people are already skeptical, if not cynical, toward the news. These interventions bet against the odds that, given the low prevalence of misinformation, true positives (rejecting or not sharing misinformation) will outnumber false positives (rejecting or not sharing reliable information).

This bet has been lost in some experimental settings (Clayton et al., 2020; Hoes et al., 2024; Modirrousta-Galian and Higham, 2023) and suggests that the risk of false positives should be systematically investigated. The most basic way to test it is to include true items in addition to false items (Guay et al., 2023b). A more sophisticated way is to use signal detection theory to detect shifts in conservative response bias and distinguish it from discrimination (Modirrousta-Galian and Higham, 2023). Beyond this, one can approximate effectiveness outside experimental settings by adjusting effect sizes to how often people encounter the test items in daily life. Suppose an intervention raises skepticism by 0.2 for true news and 0.6 for false news; the net discernment is 0.4 ($0.6 - 0.2$). Outside the lab, where true news is more common, weighting by exposure can flip the sign; with a 4:1 true to false ratio ($0.6 - (4 \times 0.2) = -0.2$). This illustrates a simple prevalence-weighted discernment, akin to Guess et al. (2020a). A stronger approach is to weigh at the item-level (Allen et al., 2024) and at the individual-level.

People are overly skeptical of reliable information

In most countries, trust in the news is low and steadily declining (Fletcher et al., 2025). Worldwide news avoidance is rising, in 2017 29% reported that they sometimes or often avoid the news, up to 39% in 2024 and 40% in 2025 (Newman et al., 2024). Interest in the news is falling: while 63% of people said they were very or extremely interested in the news in 2017, that number dropped to 46% in 2024 (Newman et al., 2024). Engagement with the news, both online and offline, is also in slow but steady decline (Altay et al., 2024b). People browse social media with a “general skepticism mind-set” and report not trusting most of the news they encounter on these platforms (Fletcher and Nielsen, 2019). On average, people in the United States and Europe can distinguish reliable news outlets from unreliable ones: when asked to rate the reliability of news outlets, their ratings strongly correlate with those of professional fact-checkers (Pennycook and Rand, 2019a; Schulz et al., 2020). In the United States, this correlation is mainly driven by their distrust of hyperpartisan and fake news websites, not their trust in mainstream media, which fact-checkers trust much more than laypeople do (Pennycook and Rand, 2019a). In experiments, participants rate false headlines as much less accurate than true headlines but are

more likely to incorrectly rate a true headline as false than to rate a false headline as true (Batailler et al., 2022; Bryanov and Vziatysheva, 2021; Pfänder and Altay, 2025). This is well encapsulated in the words of this participant in the Jahanbakhsh et al. (2021) study, who justified rating a claim as false by writing, “I’d rather not believe something that is true than believe something that is actually false” (p. 12). This observation reflects broader findings from psychology showing that people rely excessively on their prior knowledge, beliefs, and experiences, as opposed to external information (Bailey et al., 2023). In various advice-taking and social learning tasks, people consistently underuse social information and thereby “waste information” (Morin et al., 2021). This tendency to prioritize one’s own perspective over external insights, known as “egocentric discounting,” hinders learning and decision-making. This skepticism leads people to dismiss information that contradicts their priors (Taber and Lodge, 2006) and to reject reliable information more often than they accept false information (Mercier, 2020).

These high levels of skepticism suggest that interventions aimed at reducing trust in unreliable sources are likely to encounter floor effects. Most people already distrust sources that fact-checkers consider unreliable, so there is little room for improvement. Targeting the minority who consume news from these unreliable sources will help, but if they consume and share misinformation for noninformational purposes, the accuracy of the source may be of secondary importance to them. More broadly, this suggests promoting trust in reliable sources may be more effective than encouraging people to be more “critical” than they already are (Acerbi et al., 2022; boyd, 2017; Mihailidis, 2008). Instead of training laypeople to behave like fact-checkers by doing their own research and verifying the information they encounter (things that people who engage with misinformation do too much of; Marwick and Partin, 2022), it might be more fruitful to help them identify reliable sources of information and give them reasons to trust these sources as much as fact-checkers do.

Misinformation is a symptom of deeper problems

Unreliable news websites that frequently publish and share misinformation position and define themselves in opposition to the establishment, claiming to reveal truths that mainstream media are hiding (Holt et al., 2019). This type of counter-discourse does not appeal to most people, but it resonates with those who have low trust in institutions or high partisan animosity (Drażkiewicz, 2022; Marwick and Partin, 2022; Zimmermann and Kohring, 2020). For instance, one reason people believe in conspiracy theories is that their anti-elite attitudes and distrust of institutions makes these narratives appealing (Freeman et al., 2022). In general, the acceptance of information is strongly influenced by its alignment with prior beliefs, attitudes, and values (Nickerson, 1998). Similarly, people are not randomly exposed to information but self-select into information environments based, in part, on congruence with their priors (Arceneaux and Johnson, 2013). For example, political misinformation consumption is predominantly politically congenial (Guess et al., 2020b), and people who share political misinformation are animated by partisan motives, notably to derogate political opponents (Osmundsen et al., 2021). In the United States, misinformation consumption is heavily concentrated among older Republicans (Guess et al., 2021a), partly due to their higher levels of affective polarization and their appetite for hyperpartisan news (Lyons et al., 2024).

Among other factors, low trust in institutions and affective polarization create fertile ground for misinformation to take hold. Rumors and misinformation naturally emerge and thrive in such environments. Interventions against misinformation are doomed to play “whack-a-mole”: as soon as one piece of misinformation is debunked, another appears elsewhere. If there is a demand for misinformation, the offer will meet it. When misinformation is used to justify preexisting beliefs and attitudes and to fulfill social goals such as expressing dissatisfaction with political institutions, debunking can only do so much. And if older people are indeed responsible for most of the misinformation consumed and shared online, interventions should target them and not focus exclusively on young people.

Consequential misinformation comes from the top

Not all posts have equal impact. Misinformation shared by the average social media user has little reach or impact compared to that shared by politicians with millions of followers (Bennett and Livingston, 2020; Nyhan, 2021). In the United States, a small number of actors gave COVID-19 misinformation, with most of its visibility on social media (CCDH, 2021). Mainstream media inadvertently (and sometimes purposefully) amplify misinformation, and it has been suggested that, given their reach and credibility, mainstream media coverage is the greatest source of misinformation (Tsftati et al., 2020; albeit empirical support is scarce (Altay et al., 2023c, 2025). The role of social media influencers has been widely discussed, with some predicting that the 2024 US presidential election would be the “Influencer election” (Miller, 2024). However, it remains unclear to what extent they spread misinformation compared with other similarly large sources of information and how receptive their audiences are to it (Klüser and Hoes, 2024).

Misinformation from the top not only risks misleading people, but it can also distort social norms and alter what is considered socially or morally acceptable. For instance, when high-profile leaders, such as the US president, share xenophobic misinformation, they signal to the public that such views are legitimate and acceptable to express in public. This emboldens individuals to express harmful stereotypes more openly, which can lead to increased prejudice, discrimination, or even aggression toward marginalized groups.

Because most interventions target laypeople, they inevitably overlook the most consequential misinformation. Interventions aimed at the key actors—such as certain politicians, entrepreneurs, or influencers—who give misinformation its visibility would be far more effective than those currently in place. Preventing an individual with 1 million followers from sharing misinformation would be as effective in reducing the visibility of misinformation as preventing 5000 median social media users with 200 followers from doing so. Laypeople are certainly part of the problem, and some interventions should target them, but more efforts should go into designing interventions directed at powerful actors rather than ordinary users.

False information is not necessarily harmful

Misinformation is most often defined in terms of veracity, and the broader fight against misinformation is motivated by a willingness to reduce false beliefs, from which

detrimental attitudes or behaviors may arise. Yet, false information is not necessarily harmful, and true information can be used in misleading and harmful ways (Søe, 2021). For instance, ironic or satirical statements are literally false but are not necessarily misleading, as they can imply something that is not false. More broadly, information need not be false to be problematic; hateful, toxic, or racist content is harmful regardless of its truthfulness or persuasiveness. The same applies to disinformation: many campaigns do not rely on false information but instead flood the zone with true, yet biased and cherry-picked, information supporting a specific narrative.

Interventions would benefit from defining misinformation in terms of misleadingness rather than veracity (Søe, 2021) and focusing on potentially harmful misleading information, regardless of its veracity. Otherwise, they risk overlooking true information used in harmful ways or paying too much attention to benign false information. Focusing on harmful or misleading information could increase the impact of interventions, but it also creates new challenges. It is easier to fact-check blatantly false statements than it is to fact-check true statements used in misleading ways. Conceptually, broader definitions of misinformation risk becoming too vague for empirical purposes and overly subjective. This problem will be further exacerbated by negative perceptions of fact-checking—particularly among those on the political right and individuals with anti-elite worldviews (Saltz et al., 2021)—as well as by common misunderstandings regarding interventions against misinformation (Lyons et al., 2020). Since some people are already doubtful about clear-cut cases of fact-checking, more complicated and subtle fact-checks are likely to face even stronger criticism and attacks.

Pragmatic definitions of misinformation focusing on subtle properties like misleadingness need to account for the context in which a statement was uttered and what the speaker *meant* (not only what they *said*). This raises fundamental challenges for the automation of misinformation detection (Søe, 2021), with even the best models struggling to account for context (Narayanan and Kapoor, 2024), and it also calls into question our ability, as humans, to infer people's mental states based on the digital traces they leave (boyd and Crawford, 2012).

Since outright false content is relatively rare and harmful content often takes subtler forms that do not depend on outright falsehoods, it may not be worthwhile to focus too heavily on “fake news” or on content that fits narrow definitions of misinformation (such as fact-checked false news). Instead, efforts should be directed toward improving the broader information environment and mitigating the impact of more subtle or insidious forms of misinformation and harmful content (Allen et al., 2025; Altay et al., 2023a).

One way forward is to target content most likely to generate harmful misperceptions based on its reach (Allen et al., 2024). For instance, rather than debunking false claims from low-reach sources that audiences find implausible, it may be more effective to address misleading or cherry-picked information disseminated by influential accounts with greater visibility and credibility. Ideally, the effect of such content would be tested in a sample of the population to assess whether it produces misperceptions, after which additional context or a correction could be added to mitigate its effects.

Table 1 summarizes the five arguments presented in Section “Five arguments about online misinformation,” as well as their rationale and implications for interventions against misinformation.

Table 1. The main arguments are laid out in column 1, the rationale behind the arguments in column 2, and the implications of the arguments for interventions against misinformation in column 3.

Arguments	Rationale	Implications
1. Exposure to misinformation is low.	(a) Unreliable news represents 5% of people’s news diet, and a small group of active, vocal users consume most of it. (b) Most people consume little news and show limited interest in politics.	(A) Interventions against misinformation are bound to have small effects, especially compared with interventions that increase the uptake of reliable information. (B) Many people hold false beliefs because they are uninformed rather than misinformed.
2. People are overly skeptical.	(c) People distrust unreliable news outlets and do not trust reliable news outlets enough. (d) People underuse social information and are more likely to reject reliable information than to accept misinformation.	(C) There is more room to improve trust in reliable sources than to reduce trust in unreliable ones. (D) Instead of teaching laypeople to be more critical than they already are, it is more fruitful to foster trust in reliable sources.
3. Misinformation is a symptom of deeper problems.	(e) Counter-discourses based on misinformation attract people with low trust in institutions or strong partisan animosity. (f) Misinformation is more prevalent in countries with weaker democratic institutions and lower press freedom.	(E) Misinformation will persist as long as people distrust institutions and remain affectively polarized. (F) Strengthening democratic institutions and rebuilding trust in them could help address the conditions that make misinformation appealing.
4. Consequential misinformation comes from the top.	(g) A small number of powerful individuals give misinformation most of its visibility.	(G) Efforts to combat misinformation should not rest primarily on laypeople’s shoulders and instead need to target super-spreaders.
5. False information is not necessarily harmful.	(h) False information can be used in nonmisleading ways, whereas true information can be used in misleading ways.	(H) Interventions should prioritize harmful, deceptive, or misleading content over information that is merely false.

Current interventions against misinformation

This section reviews the most common interventions against misinformation, highlights their strengths and limitations, and discusses their effectiveness outside controlled experimental settings.

Fact-checking

A large body of work has shown that fact-checks are effective in reducing misperceptions and rarely backfire (Walter and Murphy, 2018; Wood and Porter, 2019). If fact-checks can correct beliefs in the short run (Carey et al., 2022), they largely fail to affect

(a) beliefs when they run counter to political elites' cues (Berlinski et al., 2021), (b) attitudes such as preferences for a political candidate or vaccine attitudes (Barrera et al., 2020; Porter et al., 2023), and (c) behaviors such as voting or vaccination intentions (Nyhan et al., 2019).

Moreover, fact-checkers cannot keep pace with the production and spread of misinformation. Crowd-sourced fact-checking is a promising solution. Research shows that politically diverse crowds can reliably identify misinformation at scale when enough people participate (Allen et al., 2021a; Martel et al., 2024).

Because most studies on fact-checking have been conducted in controlled experimental settings, its real-world *effectiveness* remains uncertain. First, can fact-checks reach individuals who have been exposed to false claims and who believe them? Empirical evidence regarding the reach of fact-checking is scarce and not optimistic. Guess et al. (2020b) found that only ~3% of people who read an article from an untrustworthy website also read a fact-check of the article. Second, can fact-checks convince the misinformed, or do they only preach to the choir? Mainstream media, which performs most fact-checking, is distrusted by heavy false news consumers. The best evidence suggests that although fact-checks are less persuasive for those with low trust in the news, they are nonetheless effective (Martel and Rand, 2024).

More broadly, it is unclear whether the most misleading and harmful claims can be convincingly fact-checked (Park et al., 2021). The most problematic type of information is not blatantly false information, such as fake news, but truths used in misleading ways and the strategic omission of true information (Broockman and Kalla, 2025). For instance, exposure to partisan news may be more harmful than exposure to false information through well-documented mechanisms such as partisan coverage, agenda-setting, and framing (Weeks et al., 2021). Even if partisan media does not strongly influence people's attitudes or behaviors directly, it can have harmful indirect effects, such as fueling partisan animosity or eroding trust in mainstream media (Guess et al., 2021b).

Finally, it has been suggested that people share misinformation not because they are *unable* to evaluate its veracity, but because they are *unwilling* to do so and have other motivations than accuracy (Pennycook and Rand, 2019b). Misinformation experts identify partisanship, identity, and confirmation bias as the primary drivers of misinformation (Altay et al., 2023b)—far outweighing incidental and nondeliberate factors such as inattention, lack of education, or low literacy. Motivated reasoning is a long-known obstacle to the success of fact-checking. When a headline is shared to denigrate political opponents or express dissatisfaction with institutions, its veracity often takes a backseat to its political alignment and perceived usefulness to the political cause. Partisan dynamics in news sharing on social media extend to fact-checks themselves, which are frequently used as political weapons—shared to discredit opponents or rally support for one's side (Shin and Thorson, 2017).

Nudges

A nudge is a subtle change in the decision environment that influences behavior without restricting choices. One of the most widely used nudges against misinformation by social media companies in recent years is friction. It has been implemented in various ways, but

the core principle is to increase the number of clicks required to share content, such as by asking users whether they want to read the article before sharing it. Friction is effective at deterring sharing (Barrera et al., 2020) and requires only minimal tweaks to platform design. The main issue is that friction reduces not only the sharing of misinformation but also that of reliable information. And since people encounter reliable information more often than misinformation, friction is likely to harm sharing discernment and reduce the sharing of misinformation at the expense of reliable information. This is concerning because, unlike the negative effects of misinformation, the positive effects of access to reliable information are well documented (Snyder and Strömberg, 2010). It is crucial to develop and test forms of friction that specifically target unreliable content. However, since friction inevitably reduces overall sharing, efforts should compensate for this decrease, notably by promoting the sharing of news from reliable sources. This may be especially important for those least interested in the news, who get most of their news incidentally while scrolling social media (Fletcher and Nielsen, 2018).

In the scientific literature, the most famous nudge against misinformation is the accuracy nudge. Prompting people to think about the accuracy of headlines, such as by asking them to rate their accuracy, reduces the sharing of false news and, to a lesser degree, boosts true news sharing (Pennycook et al., 2021). This nudge is easy to implement and has proven *efficacious* across countries (Arechar et al., 2023). Yet, to be *effective* outside experimental settings, and in the long run, it needs to be paired with other interventions to minimize habituation effects (Roozenbeek and van der Linden, 2022). A recent field experiment demonstrated that the accuracy nudge, delivered through ads on social media, can deter the sharing of misinformation. However, its overall impact is minimal because most people do not share enough news content, let alone misinformation, for the nudge to be effective (Lin et al., 2024). Indeed, most social media users are “lurkers”; they consume content without producing or sharing it, and the majority rarely, if ever, post about politics or social issues (McClain, 2021). This limits the effectiveness of all interventions aimed at reducing misinformation sharing.

Finally, source labels providing information about the reliability of a source have been developed by some companies (e.g. NewsGuard) and have been embraced by some search engines (e.g. Microsoft Edge). Source labels are easy to implement, but the evidence for their effectiveness is mixed (Aslett et al., 2022). To be effective, the labels may need to provide users with good reasons to trust the ratings. Food labels like Nutri-Score are effective despite the absence of such reasons (Shangguan et al., 2019), but users may see them as more objective and thus require less explanation. Given that most people already distrust and avoid unreliable sources, labels may be particularly useful if they provide reasons for people to trust reliable sources more, such as showing the extent to which fact-checkers trust them or how frequently their articles are fact-checked as false.

Media literacy training and prebunking

Instead of correcting misperceptions after the harm is already done, and often difficult to mitigate (Lewandowsky et al., 2012), prebunking works by preventing misperceptions from taking hold in the first place. A growing body of research has shown that various forms of prebunking (e.g. fact- or logic-based, gamified or not) can help people identify

manipulation techniques (Roozenbeek et al., 2020). More traditional educational interventions, such as media or digital literacy programs developing critical thinking skills, can also help people detect misinformation (Guesset al., 2020a), but these interventions are not always effective, even when they last up to an hour (Badrinathan, 2021).

The goal of most of these interventions is to foster skepticism toward unreliable sources and false information. Considering the low prevalence of misinformation and people's distrust of unreliable sources, these interventions are bound to have small effects outside experimental settings. But some could even have deleterious effects by increasing skepticism of both unreliable and reliable sources (Guess et al., 2020a; boyd, 2017; Clayton et al., 2020; Mihailidis, 2008; Modirrousta-Galian and Higham, 2023; Vraga et al., 2021). For instance, Clayton et al. (2020) found that a general warning highlighting the existence of misleading information and of malicious actors increased skepticism toward reliable news. The last sentence of the warning used in Clayton et al. (2020) reads "It is important to remain skeptical when reading headlines [. . .]." Similarly, Facebook tips to spot false news, which were promoted at the top of users' news feeds in 14 countries in 2017, aimed at increasing skepticism. The first Facebook tip was "Be skeptical of headlines," and the last tip was "Some stories are intentionally false." In experimental settings, these tips have been shown to "reduce the perceived accuracy of both mainstream and false news headlines" (Guess et al., 2020a: 1). The tips exemplify the premises of these interventions: they treat gullibility as the main problem and assume accordingly that the solution is to cultivate skepticism. Similarly, interventions based on inoculation theory aim explicitly at building resistance to persuasion and attitude change (Compton et al., 2021). However, there are reasons to think that people are excessively skeptical (Fletcher and Nielsen, 2019; Mercier, 2020; Morin et al., 2021) and that changing people's minds is a daunting task—it requires trust and good arguments (Coppock, 2023; Mercier, 2020). Instead of merely encouraging skepticism, digital literacy and prebunking interventions could aim to cultivate trust in reliable sources and encourage individuals to revise their beliefs in the light of new evidence. Recent evidence shows that media literacy tips promoting reliable news can improve discernment and trust in the traditional news media (Altay et al., 2024a), which suggests that trust-enhancing tips should at minimum be combined with widely used skepticism enhancing tips.

Another premise of these interventions is that people share and believe misinformation because they *lack* something, whether it is critical thinking, attention, or skills. There is some truth to this, but it misses the bigger picture: people are often *motivated* to engage with misinformation (Lyons et al., 2024). They are not passively influenced by it, but actively look for it (Motta et al., 2024a), use it to define themselves, and proudly harbor it (Marwick and Partin, 2022). People do not simply "fall into the rabbit hole"; they jump in and dig (Budak et al., 2024; Chen et al., 2023). Motivations to derogate political opponents and distrust of institutions will not be cured by skills and knowledge. As Bennett and Livingston (2018) wrote in a seminal article,

Part of this broadening of perspective is to resist easy efforts to make the problem go away by fact-checking initiatives and educating citizens about the perils of fake news. Many citizens actively seek such information to support identities and political activities [. . .]. (p. 135)

We have seen in previous sections that people are less gullible than commonly thought; the same goes for conspiracy believers (de Wildt and Aupers, 2023; Marwick and Partin, 2022; Mercier, 2020). They have a “chronically distrusting mind-set” (Newman et al., 2022) and are “epistemic individualists” (Tomas et al., 2022), encouraging everyone to “think for themselves” and “do their own research.” They are more likely to say that the opinion of someone has no value compared to their own experience, and that they don’t need others to understand the world. For them, knowledge ought to be constructed from the bottom up and should not be “blindly” acquired based on trust in a top-down fashion (Marwick and Partin, 2022). QAnon followers call this knowledge-making activity “baking,” and many of the guidelines structuring this activity are similar to those advocated in media literacy courses. Instead of accepting every conspiracy theory they encounter, members on the popular r/conspiracy subreddit (de Wildt and Aupers, 2023), the 8chan imageboard (Marwick and Partin, 2022), or the flat earth community (Dieguez and Wagner-Egger, 2021), actively debate the veracity of competing conspiracy theories. They often mock those who believe in what they consider “wrong” or absurd theories (such as lizard people or a domed flat earth) and, at times, suspect them of being infiltrators seeking to discredit more plausible and legitimate conspiracy theories (de Wildt and Aupers, 2023).

Future directions for interventions against misinformation

The final section builds on the five arguments presented in Table 1 and offers ways to adapt and improve interventions against misinformation.

Promote reliable information

To reduce the negative consequences of false beliefs, interventions do not need to focus solely on misinformation. Misinformation consumption is relatively low, yet false beliefs are widespread. This raises the question: Where do false beliefs come from if not exposure to misinformation? People can hold false beliefs and make suboptimal decisions, not because they have been misinformed, but simply because they are uninformed. For instance, people can be COVID-19 vaccine-hesitant because they have not been exposed to reliable information about the vaccines and are uncomfortable with needles or what they think the vaccine contains. By default, people have few reasons to trust vaccines. Their process is profoundly counterintuitive (Miton and Mercier, 2015).

Most of our intuitions about the world are at odds with scientific knowledge and medicine. When not exposed to reliable information, these intuitions lead us astray on several topics, such as the safety of vaccines, genetically modified organisms, or nuclear energy (Mercier, 2020; Mercier and Altay, 2022). Without access to scientific knowledge, one necessarily holds false beliefs about the formation of mountains, as the theory of tectonic plates is outside the scope of our intuitive understanding of geology. Since far more people are *uninformed* because they do not follow the news than *misinformed* because they consume misinformation (Allen et al., 2020), the fight against false beliefs is most likely to be effective when pursued in concert with the broader fight for reliable information (Acerbi et al., 2022).

So far, I have referred to the uninformed in broadly negative terms, but the uninformed do not necessarily hold false beliefs or engage in harmful behaviors. Most people do not know exactly how vaccines work or how their effectiveness is tested, but still get vaccinated because they trust scientists and the government. They are not considered a priority target population for interventions against misinformation because they already “do the right thing” and do not believe misinformation. Yet, they represent an unused potential for interpersonal mass persuasion (Altay and Mercier, 2020; Goldberg et al., 2019). When the uninformed encounter misinformation, they are often disarmed. Interventions aimed at strengthening their argumentative arsenal (by providing them with reliable information that contradicts misinformation) could allow them to better resist persuasion attempts and convince their peers via interpersonal communication (Altay and Mercier, 2020). The role of interpersonal communication goes beyond helping the uninformed convince a vaccine-hesitant relative at a family dinner. This could enable interventions to scale up and reach segments of the population typically resistant to them (e.g. see: “postinoculation talk”; Ivanov et al., 2012). Interpersonal communication is known to play a key role in relaying information from the media to the public (Katz and Lazarsfeld, 1955: 55). In sum, individuals who trust institutions can contribute to the dissemination of reliable information if they are supplied with it.

Increase trust in reliable sources

The division of cognitive labor in modern societies prevents individuals from independently verifying all the information they encounter. Instead, they must largely rely on others and believe by delegation, and accept information because it comes from reliable sources without being able to vet its content. To evaluate the veracity of the content they encounter online, people rely on snap judgments (e.g. the source, the font, and the quality of visuals) and are not willing to engage in laborious content verification (Metzger et al., 2010). Interventions could be more effective if they operated at the source level rather than at the content level. For instance, critical thinking programs are increasingly focusing on helping people to trust the right sources and be better social learners, rather than to believe the right content and be better individual reasoners (Pasquinelli et al., 2021). This is important given that the logical fallacies students learn about in critical thinking courses, and which are supposed to help them detect misleading arguments, can be used in nonfallacious ways (Hahn and Oaksford, 2007).

If the main epistemic flaw of people who believe in conspiracy theories is to rely excessively on their knowledge and not enough on others (Tomas et al., 2022), then promoting intellectual humility (i.e. making people aware of their ignorance) and open-minded thinking could help (i.e. increasing their ability to consider viewpoints they disagree with). For instance, people who are more open to revising their beliefs are less susceptible to misinformation (Roozenbeek et al., 2022). Similarly, people higher on intellectual humility are less likely to believe and share hostile conspiratorial news (Marie and Petersen, 2024). Yet, there is no guarantee that intellectual humility and open-mindedness can effectively be fostered, especially among people who consume misinformation. And what happens once people recognize their ignorance and become open to opposing viewpoints? Will they know who to trust? Media literacy training could

bridge this gap by teaching people how to identify reliable sources. In today's high-choice media environment, however, people can easily avoid news in favor of entertainment (Prior, 2003), just as they can circumvent misinformation interventions by using platforms with weaker content moderation. This is why media literacy should not only promote critical evaluation of information but also encourage engagement with credible sources and platforms.

Many interventions that have proven to be effective in reducing belief in misinformation could be adapted to promote trust in reliable sources. For example, instead of teaching people to recognize manipulation techniques, it may be more fruitful to explain to them the verification techniques that journalists use to fact-check information and produce high-quality information - much like interventions helping individuals better understand how news is produced, verified, and disseminated (i.e. "procedural news knowledge"; Amazeen and Bucy, 2019).

Ultimately, journalists need to step up. They need to prioritize impartiality, avoid sensationalist language, and "leave their personal feelings at the door" (Newman et al., 2025). Audiences demand accuracy and truth-telling, expecting reporters to stick to facts and avoid speculation (Newman et al., 2025). Transparency matters as well; people seek fuller source disclosure, clear funding details and conflict-of-interest statements, prominent error corrections, and clearer distinctions between news and opinion. Audiences expect better reporting: more in-depth investigations and expert coverage, fewer attention-grabbing headlines chasing clicks (Newman et al., 2025). Although meeting these expectations is no easy feat, progress on any of these fronts will help the news media gain the trust of the public.

Reduce the demand for misinformation

Misinformation consumption and sharing are unevenly distributed, with a minority of older, highly politicized users responsible for the majority of it. This has led many scholars to suggest that instead of teaching literacy skills, the focus should be on reducing the partisan animosity that motivates misinformation consumption and sharing (Bail, 2021; Lyons et al., 2024; Mihailidis and Viotty, 2017). For instance, Lyons et al. (2024) argue that "older Americans do not suffer a particular deficit in such skills and interventions aiming to improve the quality of online news sharing may be better served by instead targeting the increasing partisan animosity among these news consumers" (p. 25). Interventions to reduce partisan animosity propose not only to correct misperceptions (and metamisperceptions) about the outgroup, but also to highlight commonalities so that partisan identity encompasses the outgroup, develop dialogue skills to communicate across political divides, foster positive contact, and address the institutional factors that give rise to a polarizing environment (Hartman et al., 2022). If successful, such interventions may have a variety of positive effects that go beyond reducing belief in misinformation. For example, they could increase people's willingness to discuss with political opponents (Bail, 2021). Researchers and nonprofit organizations have tested several depolarization techniques (Hartman et al., 2022), many of which are promising, although concerns remain about their replicability and generalizability to other contexts and countries. Beyond polarization, more should be done to reduce the demand for

misinformation and the reasons why some individuals actively seek out and believe such content (Budak et al., 2024). For instance, beyond reducing the visibility of racist or sexist misinformation, we must also reduce the attitudes that sustain it and prevent these attitudes from taking hold in the first place by fostering environments that reward empathy and nuance. This is a daunting task, and there is no one-size-fits-all solution, but strengthening democratic norms and institutions should help.

Target super-spreaders

The most consequential misinformation often comes from the top. Some political elites routinely use their influence to spread falsehoods. Yet, few interventions target them and instead largely focus on ordinary users. One important exception is a study by Nyhan and Reifler (2015) testing the effectiveness of fact-checking on elites. They sent letters to US state legislators to warn them about the reputational risks of making questionable statements. Those who were sent these letters were “substantially less likely to receive a negative fact-check rating or to have their accuracy questioned publicly” (p. 1). These findings suggest that even if fact-checking has little direct effect on attitudes or behaviors in the wild, its mere existence could deter powerful actors from sharing false claims (see also, Lim, 2018). The indirect effect of fact-checking as a reputational threat deserves further investigation. For instance, could it incentivize politicians to be more accurate in political debates? And are some ways of highlighting the reputational threat of fact-checking more effective than others? More broadly, how can existing interventions against misinformation be adapted to target politicians and other super-spreaders of misinformation?

Beyond politicians, social media influencers can be super-spreaders of misinformation, but their reach and parasocial closeness can also be leveraged to disseminate reliable information. A TikTok field experiment found that providing content creators with evidence-based toolkits (and brief training) increased the proportion of evidence-based mental health content they posted (Motta et al., 2024b). A WhatsApp and Instagram field experiment showed that following the news influencer HugoDécrypte helps users be more informed (Altay et al., 2025).

An effective response to misinformation requires multilevel interventions tailored to the distinct roles actors play within the (mis)information ecosystem. For powerful actors, regulatory, institutional, and accountability measures are crucial. For laypeople, individual-level interventions can help, but more ambitious interventions that address the drivers of susceptibility to misinformation would be more effective. Future research should seek to clarify which interventions work best for which groups and how to best target misinformation super-spreaders.

Strengthening norms and institutions

Unfortunately, the reasons behind partisan animosity and low trust in institutions are complex and multifactorial, and there is no quick fix. To make things worse, low trust may often be justified and stem from systemic inequalities (Drażkiewicz, 2022). For instance, conspiracy ideation is more prevalent among individuals experiencing

financial insecurity due to objective material strain (Adam-Troian et al., 2023). Conspiracy ideation is also higher in countries with greater corruption, lower press freedom (Alper, 2022; Cordonier and Cafiero, 2024), and higher unemployment rates (Cordonier et al., 2021).

Systemic interventions, aimed at reducing inequalities, precarity, unemployment, or corruption, would be effective in addressing the broader information disorder problem because they would reduce the demand for such content (Kuo and Marwick, 2021; Marwick, 2018). Systemic interventions are more costly than psychological ones, but their implementation would bring benefits to society beyond reducing misinformation. While both types of interventions are complementary, researchers must ensure that the promotion of psychological interventions does not reduce support for systemic interventions (Chater and Loewenstein, 2022). More modest platform-level interventions—such as promoting reliable news on social media or facilitating researchers' access to social media data—are also desirable, even if they do not directly address the root causes of the problem.

Finally, the watchdog function and impartiality of the media need to be strengthened, and the media needs to better represent diverse and minority views (Jungherr and Schroeder, 2021). The digital platforms shaping our information ecosystem need to be more transparent and provide researchers from all over the world (not just the United States) with better data access. The business model of these platforms needs to be compatible with quality journalism and not hinder forms of journalism that make people more resilient to disinformation, such as public service media (Humprecht et al., 2020). In the meantime, many of the individual-level interventions reviewed here can help, but their effectiveness could be enhanced if they also sought to increase acceptance of accurate information and to address the proximate causes of the problem (such as partisan animosity or distrust in institutions)—especially when targeting its root causes is not feasible (such as corruption, precarity or dysfunctional institutions).

Conclusion

To be clear, I do not argue that efforts to curb the spread of misinformation should be reduced or that current interventions against misinformation are futile. In various experimental settings, fact-checking, nudges, and media literacy training have proven to be efficacious, especially when combined (Bode and Vraga, 2021). The effect sizes are rather small and often short-lived. Their ecological validity is limited, and their generalizability in the Global South is uncertain. Yet, I do not dispute that fact-checking, nudges, and media literacy training can help. Instead, this contribution should be read as an invitation to think differently about the problem of misinformation and its solutions. In particular, policymakers, journalists, and researchers should not consider misinformation in isolation, as a purely cognitive problem reflecting a lack of rationality, but in its broader sociopolitical context (such as existing tensions within society, low trust in institutions, or political polarization).

The main limitation of this article is that it mostly relies on evidence from Western democracies. On the one hand, some arguments made in this article weaken when applied to the Global South. Misinformation may be more prevalent in the Global South, which

could make interventions against misinformation especially impactful there. Moreover, increasing trust in the news would be futile in countries where access to high-quality news is limited or heavily restricted, and it would be counterproductive in countries with low-quality media ecosystems.

On the other hand, some arguments strengthen when applied outside Western democracies. Promoting reliable information may be particularly effective in countries with low-quality media ecosystems, where identifying credible content is more difficult and requires greater skills. Importantly, the need for systemic solutions is higher in some Global South countries, where people have good reasons to distrust institutions, believe in conspiracy theories, and adhere to counter-discourses, given, among other things, the higher levels of corruption and lower levels of press freedom (Alper, 2022).

In conclusion, I hope this article will help interventions against misinformation adapt to the growing body of knowledge on misinformation and our information ecosystem. First, misinformation consumption is minimal in Western democracies, but many people around the world avoid the news and are not interested in politics. As a result, many hold false beliefs because they are uninformed rather than misinformed, stressing the importance of promoting reliable information. Second, people are not gullible—if anything, they err on the side of skepticism. Instead of encouraging skepticism, it would be more fruitful to build trust. Third, people rarely share misinformation publicly online, and the minority of users who regularly share it have deep-seated motivations that are difficult to address. Individual-level interventions, like digital literacy programs, will help, but systemic interventions are needed to address the root causes of the problem. Fourth, harmful misinformation often comes from the top. Efforts to combat misinformation should not rest primarily on laypeople's shoulders and need to target super-spreaders. Fifth, false information is not necessarily harmful, and true information can be used in misleading and harmful ways. It is therefore paramount to tackle harmful content, regardless of its truthfulness, and not focus on blatant forms of misinformation. Overall, more effort should be devoted to promoting reliable information, building trust, cultivating interest in news and politics, increasing elites' accountability, and reducing partisan animosity.

Funding

The author received funding from the University of Zurich, as a UZH Postdoc Grant, grant no. [FK-25-078].

ORCID iD

Sacha Altay  <https://orcid.org/0000-0002-2839-7375>

References

- Acerbi A (2020) *Cultural Evolution in the Digital Age*. Oxford: Oxford University Press.
- Acerbi A, Altay S and Mercier H (2022) Research note: fighting misinformation or fighting for information? *Harvard Kennedy School (HKS) Misinformation Review* 3(1): 1–15.
- Adam-Troian J, Chayinska M, Paladino MP, et al. (2023) Of precarity and conspiracy: introducing a socio-functional model of conspiracy beliefs. *British Journal of Social Psychology* 62: 136–159.

- Allen J, Arechar AA, Pennycook G, et al. (2021a) Scaling up fact-checking using the wisdom of crowds. *Science Advances* 7: eabf4393.
- Allen J, Howland B, Mobius M, et al. (2020) Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances* 6(14): eaay3539.
- Allen J, Mobius M, Rothschild DM, et al. (2021b) Research note: examining potential bias in large-scale censored data. *Harvard Kennedy School Misinformation Review*, 26 July. Available at: <https://doi.org/10.37016/mr-2020-74>
- Allen J, Pennycook G and Rand DG (2025) Addressing misperceptions takes more than combating fake news. *Trends in Cognitive Sciences* 29(9): 779–782.
- Allen J, Watts DJ and Rand DG (2024) Quantifying the impact of misinformation and vaccine-skeptical content on Facebook. *Science* 384(6699): eadk3451.
- Alper S (2022) There are higher levels of conspiracy beliefs in more corrupt countries. *European Journal of Social Psychology*. Epub ahead of print 8 November. DOI: 10.1002/ejsp.2919.
- Altay S and Mercier H (2020) Framing messages for vaccination supporters. *Journal of Experimental Psychology: Applied* 26(4): 567–578.
- Altay S, Berriche M and Acerbi A (2023a) Misinformation on misinformation: conceptual and methodological challenges. *Social Media + Society* 9(1): 20563051221150412.
- Altay S, Berriche M, Heuer H, et al. (2023b) A survey of expert views on misinformation: definitions, determinants, solutions, and future of the field. *Harvard Kennedy School Misinformation Review* 4(4): 1–34.
- Altay S, De Angelis A and Hoes E (2024a) Media literacy tips promoting reliable news improve discernment and enhance trust in traditional media. *Communications Psychology* 2(1): 74.
- Altay S, Fletcher R and Nielsen RK (2024b) News participation is declining: evidence from 46 countries between 2015 and 2022. *New Media & Society*. DOI: 10.1177/14614448241247822.
- Altay S, Hoes E and Wojcieszak M (2025) Following news on social media boosts knowledge, belief accuracy and trust. *Nature Human Behaviour* 9: 1833–1842. Available at: <https://www.nature.com/articles/s41562-025-02205-6>
- Altay S, Nielsen RK and Fletcher R (2022) Quantifying the “infodemic”: people turned to trustworthy news outlets during the 2020 pandemic. *Journal of Quantitative Description: Digital Media*. Epub ahead of print 26 August. DOI: 10.51685/jqd.2022.020.
- Altay S, Nielsen RK and Fletcher R (2023c) News can help! The impact of news media and digital platforms on awareness of and belief in misinformation. *The International Journal of Press/Politics*. Epub ahead of print 6 February. DOI: 10.1177/19401612221148981.
- Amazeen MA and Bucy EP (2019) Conferring resistance to digital disinformation: the inoculating influence of procedural news knowledge. *Journal of Broadcasting & Electronic Media* 63(3): 415–432.
- Anderson C (2021) Fake news is not a virus: on platforms and their effects. *Communication Theory* 31(1): 42–61.
- Arceneaux K and Johnson M (2013) *Changing Minds or Changing Channels? Partisan News in an Age of Choice*. Chicago, IL: University of Chicago Press.
- Arechar AA, Allen J, Berinsky AJ, et al. (2023) Understanding and combatting misinformation across 16 countries on six continents. *Nature Human Behaviour* 7: 1502–1513.
- Aslett K, Guess A, Bonneau R, et al. (2022) News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions. *Science Advances* 8(18): eabl3844.
- Badrinathan S (2021) Educative interventions to combat misinformation: evidence from a field experiment in India. *American Political Science Review*. Epub ahead of print 15 June. DOI: 10.1017/S0003055421000459.
- Badrinathan S and Chauchard S (2024) “I don’t think that’s true, bro!” Social corrections of misinformation in India. *The International Journal of Press/Politics* 29(2): 394–416.

- Bail C (2021) *Breaking the Social Media Prism*. Princeton, NJ: Princeton University Press.
- Bailey PE, Leon T, Ebner NC, et al. (2023) A meta-analysis of the weight of advice in decision-making. *Current Psychology* 42(28): 24516–24541.
- Barrera O, Guriev S, Henry E, et al. (2020) Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics* 182: 104123.
- Batailler C, Brannon SM, Teas PE, et al. (2022) A signal detection approach to understanding the identification of fake news. *Perspectives on Psychological Science* 17(1): 78–98.
- Bennett L and Livingston S (2020) A brief history of the disinformation age: information wars and the decline of institutional authority. *The Disinformation Age*. Epub ahead of print 6 October. DOI: 10.1017/9781108914628.001.
- Bennett WL and Livingston S (2018) The disinformation order: disruptive communication and the decline of democratic institutions. *European Journal of Communication* 33(2): 122–139.
- Berlinski N, Doyle M, Guess A, et al. (2021) The effects of unsubstantiated claims of voter fraud on confidence in elections. *Journal of Experimental Political Science*. Epub ahead of print 28 June. DOI: 10.1017/XPS.2021.18.
- Blair RA, Gottlieb J, Nyhan B, et al. (2024) Interventions to counter misinformation: lessons from the Global North and applications to the Global South. *Current Opinion in Psychology* 55: 101732.
- Bode L and Vraga E (2021) The Swiss cheese model for mitigating online misinformation. *Bulletin of the Atomic Scientists* 77(3): 129–133.
- boyd D (2017) Did media literacy backfire? *Journal of Applied Youth Studies* 1(4): 83–89.
- boyd D and Crawford K (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662–679.
- Broockman DE and Kalla JL (2025) Consuming cross-cutting media causes learning and moderates attitudes: a field experiment with fox news viewers. *The Journal of Politics* 87(1): 246–261.
- Bryanov K and Vziatyshva V (2021) Determinants of individuals' belief in fake news: a scoping review determinants of belief in fake news. *PLoS One* 16(6): e0253717.
- Budak C, Nyhan B, Rothschild DM, et al. (2024) Misunderstanding the harms of online misinformation. *Nature* 630(8015): 45–53.
- Carey JM, Guess A, Loewen PJ, et al. (2022) The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, Great Britain and Canada. *Nature Human Behaviour* 6: 236–243.
- CCDH (2021) The disinformation dozen. Available at: <https://www.theguardian.com/world/2021/jul/17/covid-misinformation-conspiracy-theories-ccdhd-report>
- Chater N and Loewenstein G (2022) The i-frame and the s-frame: how focusing on individual-level solutions has led behavioral public policy astray. *Behavioral and Brain Sciences*. Epub ahead of print 5 September. DOI: 10.1017/S0140525X22002023.
- Chauchard S and Garimella K (2022) What circulates on partisan WhatsApp in India? Insights from an unusual dataset. *Journal of Quantitative Description: Digital Media* 2: 1–42. Available at: <https://journalqd.org/article/view/2690>
- Chen A, Nyhan B, Reifler J, et al. (2023) Subscriptions and external links help drive resentful users to alternative and extremist YouTube channels. *Science Advances* 9: eadd8080.
- Clayton K, Blair S, Busam JA, et al. (2020) Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior* 42(4): 1073–1095.
- Compton J, van der Linden S, Cook J, et al. (2021) Inoculation theory in the post-truth era: extant findings and new frontiers for contested science, misinformation, and conspiracy theories. *Social and Personality Psychology Compass* 15(6): e12602.

- Coppock A (2023) *Persuasion in Parallel: How Information Changes Minds about Politics*. Chicago, IL: University of Chicago Press.
- Cordonier L and Brest A (2021) How do the French inform themselves on the Internet? Analysis of online information and disinformation behaviors. *Fondation Descartes*, 12 March. Available at: <https://hal.archives-ouvertes.fr/hal-03167734/document>
- Cordonier L and Cafiero F (2024) Public sector corruption is fertile ground for conspiracy beliefs: a comparison between 26 Western and non-Western countries. *Social Science Quarterly* 105(3): 843–861.
- Cordonier L, Cafiero F and Bronner G (2021) Why are conspiracy theories more successful in some countries than in others? An exploratory study on Internet users from 22 Western and non-Western countries. *Social Science Information* 60(3): 436–456.
- de Wildt L and Aupers S (2023) Participatory conspiracy culture: believing, doubting and playing with conspiracy theories on Reddit. *Convergence*. Epub ahead of print 2 June. DOI: 10.1177/13548565231178914.
- Dieguez S and Wagner-Egger P (2021) Réflexions sur la forme de la Terre. In: Baechler J and Bronner G (eds) *L'irrationnel Aujourd'hui*. pp. 323–400.
- Drażkiewicz E (2022) Study conspiracy theories with compassion. *Nature* 603(7903): 765–765.
- Fletcher R and Nielsen RK (2018) Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society* 20(7): 2450–2468.
- Fletcher R and Nielsen RK (2019) Generalised scepticism: how people navigate news on social media. *Information, Communication & Society* 22(12): 1751–1769.
- Fletcher R, Andi S, Badrinathan S, et al. (2025) The link between changing news use and trust: longitudinal analysis of 46 countries. *Journal of Communication* 75(1): 1–15. Available at: <https://academic.oup.com/joc/article-abstract/75/1/1/7907139>
- Freeman D, Waite F, Rosebrock L, et al. (2022) Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological Medicine* 52(2): 251–263.
- Freiling I, Krause NM, Scheufele DA, et al. (2021) Believing and sharing misinformation, fact-checks, and accurate information on social media: the role of anxiety during COVID-19. *New Media & Society*. DOI: 10.1177/14614448211011451.
- Goldberg MH, van der Linden S, Maibach E, et al. (2019) Discussing global warming leads to greater acceptance of climate science. *Proceedings of the National Academy of Sciences* 116(30): 14804–14805.
- Grinberg N, Joseph K, Friedland L, et al. (2019) Fake news on twitter during the 2016 US Presidential election. *Science* 363(6425): 374–378.
- Guay B, Berinsky A, Pennycook G, et al. (2023b) How to think about whether misinformation interventions work. *Nature Human Behaviour* 7: 1231–1233.
- Guay B, Berinsky AJ, Pennycook G, et al. (2023a) How to think about whether misinformation interventions work. *Nature Human Behaviour* 7: 1231–1233.
- Guess A, Aslett K, Tucker J, et al. (2021a) Cracking open the news feed: exploring What US Facebook users see and share with large-scale platform data. *Journal of Quantitative Description: Digital Media* 1: 1–48.
- Guess A, Barberá P, Munzert S, et al. (2021b) The consequences of online partisan media. *Proceedings of the National Academy of Sciences* 118(14): e2013464118.
- Guess A, Lerner M, Lyons B, et al. (2020a) A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences* 117(27): 15536–15545.
- Guess A, Nagler J and Tucker J (2019) Less than you think: prevalence and predictors of fake news dissemination on Facebook. *Science Advances* 5(1): eaau4586.

- Guess A, Nyhan B and Reifler J (2020b) Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour* 4(5): 472–480.
- Gwiażdźński P, Gundersen AB, Piksa M, et al. (2023) Psychological interventions countering misinformation in social media: a scoping review. *Frontiers in Psychiatry* 13: 974782.
- Hahn U and Oaksford M (2007) The rationality of informal argumentation: a Bayesian approach to reasoning fallacies. *Psychological Review* 114(3): 704–732.
- Hartman R, Blakey JW, Womick J, et al. (2022) Interventions to reduce partisan animosity. *Nature Human Behaviour* 6: 1194–1205.
- Hoes E, Aitken B, Zhang J, et al. (2024) Prominent misinformation interventions reduce misperceptions but increase scepticism. *Nature Human Behaviour* 8: 1545–1553. Available at: <https://www.nature.com/articles/s41562-024-01884-x>
- Holt K, Ustad Figenschou T and Frischlich L (2019) Key dimensions of alternative news media. *Digital Journalism* 7(7): 860–869.
- Humprecht E, Esser F and Van Aelst P (2020) Resilience to online disinformation: a framework for cross-national comparative research. *The International Journal of Press/Politics* 25(3): 493–516.
- Ivanov B, Miller CH, Compton J, et al. (2012) Effects of postinoculation talk on resistance to influence. *Journal of Communication* 62(4): 701–718.
- Jahanbakhsh F, Zhang AX, Berinsky AJ, et al. (2021) Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW1): 1–42.
- Jungherr A and Schroeder R (2021) Disinformation and the structural transformations of the public arena: addressing the actual challenges to democracy. *Social Media + Society* 7(1): 2056305121988928.
- Katz E and Lazarsfeld PF (1955) *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: Free Press.
- Klüser KJ and Hoes E (2024) Disinformation for hire? *Dispelling exaggerated concerns about social media influencers' role in spreading misinformation*. Available at: <https://osf.io/preprints/psyarxiv/axszp>
- Kuo R and Marwick A (2021) Critical disinformation studies: history, power, and politics. *Harvard Kennedy School Misinformation Review* 2(4): 1–11.
- Lewandowsky S, Ecker UK, Seifert CM, et al. (2012) Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest* 13(3): 106–131.
- Lim C (2018) Can fact-checking prevent politicians from lying? Available at: <https://www.semanticscholar.org/paper/Can-Fact-checking-Prevent-Politicians-from-Lying-Lim/81af918d88e9573aa49ed6e85109999e52a3615d>.
- Lin H, Garro H, Wernerfelt N, et al. (2024) Reducing misinformation sharing at scale using digital accuracy prompt ads. *PsyArXiv* [preprint]. DOI: 10.31234/osf.io/u8anb.
- Lyons B, Mérola V, Reifler J, et al. (2020) How politics shape views toward fact-checking: evidence from six European countries. *The International Journal of Press/Politics* 25(3): 469–492.
- Lyons B, Montgomery J and Reifler J (2024) Partisanship and older Americans' engagement with dubious political news. *Public Opinion Quarterly* 88(3): 962–990.
- Marie A and Petersen MB (2024) Moralization of rationality can stimulate sharing of hostile and false news on social media, but intellectual humility inhibits it. *Political Communication*. DOI: 10.1080/10584609.2024.2363542.
- Martel C and Rand DG (2024) Fact-checker warning labels are effective even for those who distrust fact-checkers. *Nature Human Behaviour* 8(10): 1957–1967.
- Martel C, Allen J, Pennycook G, et al. (2024) Crowds can effectively identify misinformation at scale. *Perspectives on Psychological Science* 19(2): 477–488.

- Marwick A and Partin WC (2022) Constructing alternative facts: populist expertise and the QAnon conspiracy. *New Media & Society*. Epub ahead of print 1 May. DOI: 10.1177/14614448221090201.
- Marwick AE (2018) Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technology Review* 2(2): 474–512.
- Mcclain C (2021) 70% of U.S. social media users never or rarely post or share about political, social issues. Available at: <https://www.pewresearch.org/fact-tank/2021/05/04/70-of-u-s-social-media-users-never-or-rarely-post-or-share-about-political-social-issues/>
- Mercier H (2020) *Not Born Yesterday: The Science of Who We Trust and What We Believe*. Princeton, NJ: Princeton University Press.
- Mercier H and Altay S (2022) Do cultural misbeliefs cause costly behavior? In: Musolino J, Hemmer P and Sommer J (eds) *The Science of Beliefs*. Cambridge: Cambridge University Press, pp. 193–208.
- Metzger MJ, Flanagin AJ and Medders RB (2010) Social and heuristic approaches to credibility evaluation online. *Journal of Communication* 60(3): 413–439.
- Mihailidis P (2008) *Beyond Cynicism: How Media Literacy Can Make Students More Engaged Citizens*. College Park, MD: University of Maryland.
- Mihailidis P and Viotty S (2017) Spreadable spectacle in digital culture: civic expression, fake news, and the role of media literacies in “post-fact” society. *American Behavioral Scientist* 61(4): 441–454.
- Miller G (2024) *If 2024 Is The Influencer Election, Where’s The FEC?* Austin, TX: Tech Policy Press.
- Miton H and Mercier H (2015) Cognitive obstacles to pro-vaccination beliefs. *Trends in Cognitive Sciences* 19(11): 633–636.
- Modirrousta-Galian A and Higham PA (2023) Gamified inoculation interventions do not improve discrimination between true and fake news: reanalyzing existing research with receiver operating characteristic analysis. *Journal of Experimental Psychology: General* 152(9): 2411–2437.
- Morin O, Jacquet PO, Vaesen K, et al. (2021) Social information use and social information waste. *Philosophical Transactions of the Royal Society B* 376(1828): 20200052.
- Motta M, Hwang J and Stecula D (2024a) What goes down must come up? Misinformation search behavior during an unplanned Facebook outage. *Health Communication* 39(10): 2041–2052.
- Motta M, Liu Y and Yarnell A (2024b) “Influencing the influencers”: a field experimental approach to promoting effective mental health communication on TikTok. *Scientific Reports* 14: 5864.
- Narayanan A and Kapoor S (2024) *AI Snake Oil: What Artificial Intelligence Can Do, What It Can’t, and How to Tell the Difference*. Princeton, NJ: Princeton University Press.
- Newman D, Lewandowsky S and Mayo R (2022) Believing in nothing and believing in everything: the underlying cognitive paradox of anti-COVID-19 vaccine attitudes. *Personality and Individual Differences* 189: 111522.
- Newman N, Fletcher R, Craig TR, et al. (2024) Reuters institute digital news report 2024. Available at: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- Newman N, Ross Arguedas A, Robertson CT, et al. (2025) *Reuters Institute Digital News Report 2025*. Oxford: Reuters Institute for the Study of Journalism.
- Nickerson RS (1998) Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology* 2: 175–220.
- Nyhan B (2020) Facts and myths about misperceptions. *Journal of Economic Perspectives* 34(3): 220–236.
- Nyhan B (2021) Why the backfire effect does not explain the durability of political misperceptions. *Proceedings of the National Academy of Sciences* 118(15): e1912440117.

- Nyhan B and Reifler J (2015) The effect of fact-checking on elites: a field experiment on US state legislators. *American Journal of Political Science* 59(3): 628–640.
- Nyhan B, Porter E, Reifler J, et al. (2019) Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior*. DOI: 10.1007/s11109-019-09528-x.
- Nyhan B, Settle J, Thorson E, et al. (2023) Like-minded sources on Facebook are prevalent but not polarizing. *Nature* 620(7972): 137–144.
- Osmundsen M, Bor A, Vahlstrup PB, et al. (2021) Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American Political Science Review*. DOI: 10.1017/S0003055421000290.
- Park S, Park JY, Kang J, et al. (2021) The presence of unexpected biases in online fact-checking. *The Harvard Kennedy School Misinformation Review*. DOI: 10.37016/mr-2020-53.
- Pasquinelli E, Farina M, Bedel A, et al. (2021) Naturalizing critical thinking: consequences for education, blueprint for future research in cognitive science. *Mind, Brain, and Education* 15(2): 168–176.
- Pennycook G and Rand DG (2019a) Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* 116(7): 2521–2526.
- Pennycook G and Rand DG (2019b) Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188: 39–50.
- Pennycook G, Epstein Z, Mosleh M, et al. (2021) Shifting attention to accuracy can reduce misinformation online. *Nature* 592(7855): 590–595.
- Pfänder J and Altay S (2025) Spotting false news and doubting true news: a systematic review and meta-analysis of news judgements. *Nature Human Behaviour* 9: 688–699.
- Porter E, Wood TJ and Velez Y (2023) Correcting COVID-19 vaccine misinformation in ten countries. *Royal Society Open Science*. Epub ahead of print 10 March. DOI: 10.1098/rsos.221097.
- Prior M (2003) Liberated viewers, polarized voters—the implications of increased media choice for democratic politics. *The Good Society* 11: 10–16.
- Robertson CE, Del Rosario KS and Van Bavel JJ (2024) Inside the funhouse mirror factory: how social media distorts perceptions of norms. *Current Opinion in Psychology* 60: 101918.
- Roozenbeek J and van der Linden S (2022) How to combat health misinformation: a psychological approach. *American Journal of Health Promotion* 36(3): 569–575.
- Roozenbeek J, Culloty E and Suiter J (2023) Countering misinformation: evidence, knowledge gaps, and implications of current interventions. *European Psychologist*. DOI: 10.1027/1016-9040/a000492.
- Roozenbeek J, Maertens R, Herzog SM, et al. (2022) Susceptibility to misinformation is consistent across question framings and response modes and better explained by open-mindedness and partisanship than analytical thinking. *Judgment and Decision Making* 17: 547–573.
- Roozenbeek J, Remshard M and Korychenko Y (2024) Beyond the headlines: on the efficacy and effectiveness of misinformation interventions. *Advances in Psychology* 2: e24569.
- Roozenbeek J, van der Linden S and Nygren T (2020) Prebunking interventions based on the psychological theory of “inoculation” can reduce susceptibility to misinformation across cultures. *Harvard Kennedy School Misinformation Review* 1(2): 1–23.
- Saltz E, Barari S, Leibowicz C, et al. (2021) Misinformation interventions are common, divisive, and poorly understood. *Harvard Kennedy School Misinformation Review* 2(5): 1–25. DOI: 10.37016/mr-2020-81.
- Schulz A, Fletcher R and Popescu M (2020) Are news outlets viewed in the same way by experts and the public? A comparison across 23 European countries. *Reuters Institute Factsheet*, 7 May. Available at: <https://reutersinstitute.politics.ox.ac.uk/are-news-outlets-viewed-same-way-experts-and-public-comparison-across-23-european-countries>

- Shangguan S, Afshin A, Shulkin M, et al. (2019) A meta-analysis of food labeling effects on consumer diet behaviors and industry practices. *American Journal of Preventive Medicine* 56(2): 300–314. DOI: 10.1016/j.amepre.2018.09.024.
- Shin J and Thorson K (2017) Partisan selective sharing: the biased diffusion of fact-checking messages on social media. *Journal of Communication* 67(2): 233–255.
- Simon F and Camargo C (2021) Autopsy of a metaphor: the origins, use, and blind spots of the ‘infodemic.’ *New Media & Society*. DOI: 10.1177/14614448211031908.
- Singal AG, Higgins PD and Waljee AK (2014) A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology* 5(1): e45. DOI: 10.1038/ctg.2013.13.
- Snyder JM and Strömberg D (2010) Press coverage and political accountability. *Journal of Political Economy* 118(2): 355–408.
- Søe SO (2021) A unified account of information, misinformation, and disinformation. *Synthese* 198(6): 5929–5949.
- Spampatti T (2025) Truth discernment may Not help to overcome misinformation. *Nature Climate Change* 15: 1006–1009.
- Taber CS and Lodge M (2006) Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50(3): 755–769.
- Tomas F, Nera K and Schöpfer C (2022) “Think for yourself, or others will think for you”: epistemic individualism predicts conspiracist beliefs and critical thinking. *PsyArXiv* [preprint]. DOI: 10.31219/osf.io/qgtzb.
- Tsfati Y, Boomgaarden H, Strömbäck J, et al. (2020) Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis. *Annals of the International Communication Association* 44(2): 157–173.
- Vraga E, Tully M and Bode L (2021) Assessing the relative merits of news literacy and corrections in responding to misinformation on Twitter. *New Media & Society*. DOI: 10.1177/1461444821998691.
- Walter N and Murphy ST (2018) How to unring the bell: a meta-analytic approach to correction of misinformation. *Communication Monographs* 85(3): 423–441.
- Weeks BE, Menchen-Trevino E, Calabrese C, et al. (2021) Partisan media, untrustworthy news sites, and political misperceptions. *New Media & Society*. Epub ahead of print 28 July. DOI: 10.1177/14614448211033300.
- Wojcieszak M, de Leeuw S, Menchen-Trevino E, et al. (2023) No polarization from partisan news: over-time evidence from trace data. *The International Journal of Press/Politics* 28(3): 601–626.
- Wood T and Porter E (2019) The elusive backfire effect: mass attitudes’ steadfast factual adherence. *Political Behavior* 41(1): 135–163.
- Ziemer C-T and Rothmund T (2024) Psychological underpinnings of misinformation countermeasures: a systematic scoping review. *Journal of Media Psychology: Theories, Methods, and Applications* 36(6): 397–409. DOI: 10.1027/1864-1105/a000407.
- Zimmermann F and Kohring M (2020) Mistrust, disinforming news, and vote choice: a panel survey on the origins and consequences of believing disinformation in the 2017 German parliamentary election. *Political Communication* 37(2): 215–237.

Author biography

Sacha Altay is a behavioural scientist working on misinformation, social media, and the news. At the time of publishing, he’s a post-doctoral researcher at the University of Zurich within the political science department, and holds a PhD in cognitive science from the École Normale Supérieure in Paris.